REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
24-02-2014	Final Report	1-Sep-2010 - 31-Aug-2013

4. TITLE AND SUBTITLE
Statistical Inference for Cultural Consensus Theory

5a. CONTRACT NUMBER
W911NF-10-1-0212

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

6. AUTHORS
William H. Batchelder

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
University of California - Irvine
Office of Research
5171 California Ave., Suite 150
Irvine, CA 92697 -7600

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
58261-LS.26

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
In the three years of support, almost all the proposed projects were completed and are published, in press, or under review in major peer reviewed journals. In particular Cultural Consensus Theory (CCT) models have been extended and/or invented for dichotomous data, ties in a graph, ordinal data, and continuous data. All the models now allow multiple consensus truths, heterogeneous item difficulty, heterogeneous informant competence, and heterogeneous informant biases. Hierarchical Bayesian inference has been provided for all of the CCT models, and software for this inference is freely available in journal articles and web sites. Bayesian post predictive tests have been

15. SUBJECT TERMS
Cultural Consensus Theory, Hierarchical Bayesian Modeling, Dichotomous Response Data, Multiple Consensus Truths, Ordinal Data, Continuous Data,

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
UU	UU	UU	UU		William Batchelder
					19b. TELEPHONE NUMBER
					949-824-7271

Standard Form 298 (Rev 8/98)
Prescribed by ANSI Std. Z39.18

Report Title

Statistical Inference for Cultural Consensus Theory

ABSTRACT

In the three years of support, almost all the proposed projects were completed and are published, in press, or under review in major peer reviewed journals. In particular Cultural Consensus Theory (CCT) models have been extended and/or invented for dichotomous data, ties in a graph, ordinal data, and continuous data. All the models now allow multiple consensus truths, heterogeneous item difficulty, heterogeneous informant competence, and heterogeneous informant biases. Hierarchical Bayesian inference has been provided for all of the CCT models, and software for this inference is freely available in journal articles and web sites. Bayesian post predictive tests have been developed and implemented for two critical model assumptions: (1) Are there one or more than one consensus truths? , (2) Are the items heterogeneous or homogeneous in difficulty? All of the CCT models have been applied to both simulated and real data, and it has been shown that the models perform well even with a small number of informants.

Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Received	Paper
02/12/2014 14.00	Royce Anders, William H. Batchelder. Cultural consensus theory for multiple consensus truths, Journal of Mathematical Psychology, (12 2012): 452. doi:
02/12/2014 15.00	William H. Batchelder, Gregory E. Alexander. Insight Problem Solving: A Critical Examination of the Possibilityof Formal Theory, The Journal of Problem Solving, (10 2012): 56. doi:
02/12/2014 16.00	Royce Anders, William H. Batchelder. CULTURAL CONSENSUS THEORY FOR THE ORDINAL DATA CASE, Psychometrika, (12 2013): 0. doi:
02/12/2014 17.00	Zita Oracevz, Royce Anders, William H. Batchelder. HIERARCHICAL BAYESIAN MODELING FOR TEST THEORY WITHOUTAN ANSWER KEY, Psychometrika, (10 2013): 0. doi:
02/12/2014 18.00	William H. Batchelder, Gregory E. Alexander. Discrete-State Models: Comment on Pazzaglia, Dube, and Rotello (2013), Psychological Bulletin, (11 2013): 1204. doi:
08/22/2011 3.00	Verena D. Schmittmann, Conor V. Dolan, Maartje E. J. Raijmakers, William H. Batchelder. Parameter identification in multinomial processing tree models, Behavior Research Methods, (08 2010): 0. doi: 10.3758/BRM.42.3.836
08/22/2011 4.00	Jared B. Smith, William H. Batchelder. Beta-MPT: Multinomial processing tree models for addressing individual differences, Journal of Mathematical Psychology, (02 2010): 0. doi: 10.1016/j.jmp.2009.06.007

TOTAL: **7**

Number of Papers published in peer-reviewed journals:

(b) Papers published in non-peer-reviewed journals (N/A for none)

Received		Paper
08/22/2011	1.00	William H. Batchelder, Hao Wu, Jay I. Myung. On the minimum description length complexity of multinomial processing tree models, Journal of Mathematical Psychology, (06 2010): 0. doi: 10.1016/j.jmp.2010.02.001
08/22/2011	2.00	Hao Wu, William H. Batchelder, Jay I. Myung. Minimum description length model selection of multinomial processing tree models, Psychonomic Bulletin & Review, (06 2010): 0. doi: 10.3758/PBR.17.3.275
TOTAL:		**2**

Number of Papers published in non peer-reviewed journals:

(c) Presentations

Sept.1, 2010- Aug 31, 2011 6 invited and conference talks
Sept.1, 2011- Aug. 31 2012 7 invited and conference talks
Sept.1, 2012- Aug. 31,2013 7 invited or conference talks

All these talks that were not published were presented as ppt slides. They include discussions that come from the published or in press papers cited in this Final Report.

Batchelder, W. H., and Anders, R. Cultural Consensus Theory: Comparing different concepts of cultural truth. Paper presented at Annual Meeting of the Society for Mathematical Psychology. Tufts University, July 2011.

Agrawal, K. (Presenter), and Batchelder, W.H. Cultural Consensus Theory: Estimating Consensus Graphs under constraints. Paper presented at Annual Meeting of the Society for Mathematical Psychology. Tufts University, July 2011.

Batchelder, W.H. Specification of Multinomial Processing Tree Models. Presentation at the Workshop on Multinomial Processing Tree Models. Tufts University, July 2011.

Batchelder, W.H. Statistical inference for Multinomial Processing Tree Models. Presentation at the Workshop on Multinomial Processing Tree Models. Tufts University, July 2011.

Batchelder, W.H. Some Issues in the Developing a Theory of Human Problem Representation. Keynote address at Dagstahl Workshop: Computer Science and Problem Solving. Dagstahl, Germany, September 2011.

Batchelder, W.H. Cultural Consensus Theory: Detecting Experts and Their Shared Knowledge. Invited Paper presented at DIMACS Workshop on the Science of Expert Opinion. Rutgers University, October 2011.

Batchelder, W. H., and Agrawal, K. Cultural Consensus Theory: Aggregating Complete Signed Graphs Under a Balance Constraint -- Part 1. International Sunbelt Social Network Conference XXXII, Redondo Beach, California, March 2012.

Agrawal, K. (Presenter), and Batchelder, W. H. Cultural Consensus Theory: Aggregating Complete Signed Graphs Under a Balance Constraint -- Part 2. International Sunbelt Social Network Conference XXXII, Redondo Beach, California, March 2012.

Agrawal, K.(Presenter), and Batchelder, W. H. Cultural Consensus Theory: Aggregating Signed Graphs Under a Balance Constraint. Paper presented at the Annual Meeting of the Social Computing, Behavioral Cultural Modeling, and Prediction. University of Maryland, April 2012.

Batchelder, W. H. Cultural Consensus Theory: Latest Developments. Colloquium Institute for Mathematical Behavioral Sciences, University of California Irvine, May 2012.

Matzke, D. (Presenter), Dolan, C., Batchelder, W. H., and Wagenmakers, E-J. Hierarchical Multinomial Processing Tree Models for the Pair-clustering Paradigm with Heterogeneity in Participants and Items. Paper presented at the Annual Meeting of the Society for Mathematical Psychology, Columbus, Ohio, July 2012.

Anders, R. (Presenter), and Batchelder, W. H. Cultural Consensus Theory for Multiple Consensus Truths. Paper presented at the Annual Meeting of the Society for Mathematical Psychology, Columbus, Ohio, July 2012.

Oravecz, Z. (Presenter), and Batchelder, W. H. Bayesian Hierarchical Cultural Consensus Theory. Paper presented at the Annual Meeting of the Society for Mathematical Psychology, Columbus, Ohio, July 2012.

Batchelder, W. H., Anders, R., and Oracevz, Z. The Inverse Problem of Signal detection. Invited paper read at 51st Annual Edwards Bayesian Research Conference. Fullerton. CA.February 2013

Batchelder, W. H. Cultural Consensus Theory. Invited paper in Wisdom of the Crowd Conference. Institute for Mathematical Behavioral Sciences, University of California Irvine, April 2013.

Batchelder, W. H. Cultural Consensus Theory, Invited paper in Festschrift for Professor James T. Townsend. Indiana University, April 2013.

Batchelder, W. H. Cultural Consensus Theory: The General Condorcet Model. Paper presented in invited session on Cultural Consensus theory at Annual Classification Society of North America, Milwaukee. June 2013.

Batchelder, W. H. History of Mathematical Psychology. Invited paper presented at Advances in Behavioral Sciences in the Internet Age—Selected Topics in Mathematical Psychology. Forum held at Central China Normal University, Wuhan, China. June 2013.

Batchelder, W. H. Cognitive Psychometrics. Invited paper presented at Advances in Behavioral Sciences in the Internet Age—Selected Topics in Mathematical Psychology. Forum held at Central China Normal University, Wuhan, China. June 2013.

Anders, R. (Presenter), and Batchelder, W. H. A Cultural Consensus Theory Model for the Polytomous Data Case. Paper read at the Annual Meeting of the Society for Mathematical Psychology. University of Potsdam, Potsdam, Gernany, August, 2013

Number of Presentations: 29.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received	Paper

TOTAL:

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received		Paper
02/12/2014	19.00	Stephen L. France, Mahyer Vagheti, William H. Batchelder. FlexCCT: Software for Continuous CCT, Proceedings of the 6th International Conference on Educational Data Mining. 06-JUL-13, . : ,
02/12/2014	20.00	Kalin Agrawal, William H. Batchelder. Cultural Consensus Theory: Aggregating SignedGraphs Under a Balance Constraint, Social Computing, Behavioral-Cultural Modeling, and Prediction. 03-APR-12, . : ,
08/22/2011	5.00	Giorgio Gosti and William H. Batchelder. Naming on a Directed Graph, Social Computing, Behavioral-cultural Modeling and Prediction. 29-MAR-11, . : ,
08/31/2012	7.00	Kalin Agrawal, William H. Batchelder. Cultural Consensus Theory: Aggregating SignedGraphs under a Balance Constraint, Social Computing, Behavioral Cultural Modeling, and Prediction. 04-APR-12, . : ,

TOTAL: **4**

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts):

(d) Manuscripts

Received		Paper
02/12/2014	21.00	Stephen France, William H. Batchelder. Maximum Likelihood Item Easiness Models for Test Theory Without An Answer Key, Educational and Psychological Measurement (07 2013)
02/12/2014	22.00	Royce Anders, William H. Batchelder. Cultural Consensus Theory Application Software,CCTpack, Journal of Statistical Software (12 2013)
02/12/2014	23.00	Royce Anders, Zita Oravecz, William H. Batchelder. Cultural Consensus Theory for Continuous ResponsesA Latent Appraisal Model for Information Pooling, Journal of Mathematical Psychology (12 2013)
02/13/2014	24.00	Zita Oravecz, Katherine Faust, William H. Batchelder. An extended Cultural Consensus Theory model to account forcognitive processes in decision making in social surveys, Sociological Methodology (07 2013)
02/13/2014	25.00	Zita Oravecz, Katherine Faust, Daniel A. Levitis, William H. Batchelder. Studying the existence and attributes of consensus in denitions ofpsychological concepts by a cognitive psychometric model, American Journal of Psychology (07 2013)
08/23/2011	6.00	William H. Batchelder, Royce Anders. Cultural Consensus Theory: Comparing Different Concepts of Cultural Truth, Journal of Mathematical Psychology (08 2011)
08/31/2012	8.00	William H. Batchelder, Royce Anders. Cultural Consensus Theory: Comparing different concepts of cultural truth, Journal of Mathematical Psychology (07 2012)
08/31/2012	9.00	William H. Batchelder, Gregory Alexander. Insight Problem Solving: A Critical Examination of thePossibility of Formal Theory, Journal of Problem Solving (03 2012)
08/31/2012	10.00	Zita Oravecz, Joachim Vandekerckhove, William H. Batchelder. Bayesian Cultural Consensus Theory, Field Methods (07 2012)
08/31/2012	11.00	Zita Oravecz, Joachim Vandekerckhove, William H. Batchelder. User's guide to Bayesian Cultural Consensus Toolbox, Field Methods (07 2012)
08/31/2012	12.00	William H. Batchelder. Discrete State Models of Information Processing, INTERNATIONAL Encyclopedia of Social and Behavioral Sciences (03 2012)
08/31/2012	13.00	Dora Matzke, Conor Dolan, William H. batchelder, Eric-Jan Wagenmakers. Bayesian Estimation of Multinomial Processing Tree Models with Heterogeneity inParticipants and Items, Psychometrika (07 2012)

TOTAL: **12**

Number of Manuscripts:

Books

Received	Paper

TOTAL:

Patents Submitted

Patents Awarded

Awards

Received Award 2012-2013 and 2013-2014 from the Oak Ridge Institute for Science and Education (ORISE) to work with Dr. Mike Young, a researcher at Wright Patterson Air Force Base.

Graduate Students

NAME	PERCENT_SUPPORTED	Discipline
Giorgio Gosti	0.00	
Royce Anders	0.00	
Kalin Agrawal	0.05	
Gregory E. Alexander	0.05	
James Negen	0.00	
William Leibzon	0.00	
FTE Equivalent:	**0.10**	
Total Number:	**6**	

Names of Post Doctorates

NAME	PERCENT_SUPPORTED
Dr. Zita Oravecz	1.00
FTE Equivalent:	**1.00**
Total Number:	**1**

Names of Faculty Supported

NAME	PERCENT_SUPPORTED	National Academy Member
Professor E. J. Wagenmakers	0.00	
Professor Stephen France	0.00	
Professor Joachim Vandekerchhov	0.00	
Professor Katherine Faust	0.00	
Professor Daniel A. Levitis	0.00	
Professor Emeritus A. Kimball Ror	0.00	Yes
FTE Equivalent:	**0.00**	
Total Number:	**6**	

Names of Under Graduate students supported

NAME	PERCENT_SUPPORTED	Discipline
Francis Lee	0.00	Cognitive Sciences
Linda Vuong	0.00	Economics
Gayane Gyulakopyan	0.00	Cognitive Sciences
So0bin Kristen Ahn	0.00	Cognitive Sciences
So Young Choi	0.00	Cognitive Sciences
FTE Equivalent:	**0.00**	
Total Number:	**5**	

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period: 5.00

The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:...... 5.00

The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:...... 5.00

Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):...... 5.00

Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense 0.00

The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields:...... 3.00

Names of Personnel receiving masters degrees

NAME	
Royce Anders	
Total Number:	**1**

Names of personnel receiving PHDs

NAME	
Dr. Giorgio Gosti	
Dr. Royce Anders	
Total Number:	**2**

Names of other research staff

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Sub Contractors (DD882)

Inventions (DD882)

Scientific Progress

The Scientific Progress and Accomplishments is presented in an attachment

Technology Transfer

Scientific Progress Report

During the three years of the grant, almost all the areas of the original proposed research have led to productive results, and there have been several additional related developments. The organization of the report will consist of six sections. Before detailing the progress under the grant, Section I will provide a brief review of the essential aspects of Cultural Consensus Theory (CCT). Each of the remaining sections (with the exception of Section V) will cover a major component of the proposed research. Section II will describe new software that has been developed for CCT models, the main focus of the grant. Section III will describe new CCT models that were invented in the period of the grant, and Section IV will describe tests of the one-cultural truth assumption and its extension to the case of multiple consensus truths. Section V will describe briefly other research supported by the grant, and Section VI will be a conclusion. The published work will be cited by paper numbers used in the section for downloaded papers in the final report.

I. Brief Review of Concepts in Cultural Consensus Theory

Cultural Consensus Theory is a statistical modeling approach to information pooling developed by William H. Batchelder, A. Kimball Romney, Susan Weller, and colleagues. It first appeared in the mid 1980s in a series of papers (Batchelder and Romney, 1986, 1988, 1989; Romney, Weller, and Batchelder, 1986; Romney, Batchelder, and Weller, 1987). Since its inception, CCT has become a leading methodology in cultural anthropology, e.g. Romney and Batchelder (1999), Weller (2007), and in the 1990s and early 2000s a few additional papers on CCT were published, some applying it in other areas of the social sciences. However, in the last five years, there have been major developments in CCT models supported by grants from the AFOSR and the ARO. In particular, during the period of the three-year grant from the ARO, CCT has seen applications in quantitative journals in social networks, sociology, psychology, education, and statistics (described later).

CCT consists of a collection of cognitively motivated, parametric statistical models for pooling the responses of "informants" (experts, respondents) to a series of questions about some domain of their shared knowledge or beliefs. CCT domains include areas such as eyewitness reports, forecasts of probabilistic events, folk medical beliefs, grammaticality judgments, moral or religious beliefs, scientific knowledge in various populations, judges grading student essays or sports performance, and ties in a social or covert network. CCT is intended to operate with relatively small groups of heterogeneous informants, and it applies in situations where researchers know how to pose questions in some response format to the informants, but they do not know the 'consensus answers,' if any, a priori. The data structure of CCT requires that each informant answer each of a fixed series of questions (procedures have been developed for dealing with missing data as well).

The overall aims of a CCT model can be described in general terms as follows. Let X_{ik} be a random variable representing the response of informant *i* to question *k*, and $\mathbf{X} = (X_{ik})_{N \times M}$ the random *response profile matrix* for *N* informants each answering *M*

questions. The goal of applying a CCT model to such data includes the following: (a) Identify one or more latent “cultural groups” of informant that share consensus answers to the questions; (b) Decide if the statistical model used to do this is sufficiently supported by the data, (c) If statistical support for the model is found, then the goal is to estimate the parameters of the model. Most CCT models have parameters for the consensus answers to the questions, the “cultural salience” (difficulty) of each question, the “cultural competence” (calibration) of each informant, the “response biases” of each informant. If more than one consensus truth is estimated, corresponding group membership parameters for each informant are estimated. It is important to note that all CCT inference is based only on the obtained data **X,** and no exogenous information is used in the estimation.

At the time of the proposal, the only generally available statistical software for a CCT model was due to Stephen P. Borgatti (1900, 1992). This software was for an early CCT model called the General Condorcet Model (GCM, Batchelder and Romney, 1988) for dichotomous (True/False) questionnaire data (it could also be applied to multiple choice data with additional assumptions). Also the model assumed only one consensus answer key, and it assumed that all items were equally difficult and that all informants had neutral (50:50) guessing biases. While many published articles in cultural anthropology made use of this software (e.g. Weller, 2007), it was obvious that there was a need for models to be developed for other testing formats, e.g. ordinal (Likert) data, continuous data; and in each case it was important to have available user-friendly software for the new models. Some effort in this direction was provided for the GCM in Karabatsos and Batchelder (2003), but the inference software was not presented in a way that unsophisticated users could employ it, and as a consequence it had almost no effect on the practice of users of CCT. These limitations were the main basis of what motivated this grant that was funded by the Army Research Office.

II. New Statistical Software

The original proposal stated the following: “Despite its popularity, CCT is handicapped by the availability of very limited software that can handle only special cases of the theory. The principle objective of the proposal is to develop a user-friendly and freely available software package for cultural consensus theory (CCT) based on modern computational Bayesian approaches to statistical inference. “

There have been three major developments supported by the ARO grant in this area. First, in two papers to appear in *Field Methods* (Papers 10, 11), the weaknesses in the previous software for the GCM have been overcome. Paper (10) provided the Bayesian theory and the basics of a Graphic Users Interface (GUI) for the GCM, and paper (11) was a User’s Guide for the GUI. *Field Methods*, a leading journal for cultural anthropology studies, accepted this paper in 2012; however, the large backlog in that journal prevented its publication until now (available now On-line, *Field Methods*, February 10, 2014). However, the two papers have been available on Dr. Zita Oracevz’s website (bayesian.zitaoravecz.net) since 2012, and she offers help to people trying to use the GUI (Zita is the Post Doctoral Fellow supported by the grant). The style of the paper and the GUI is to make it understandable and usable by researchers with limited statistical and

computational backgrounds. This software improves upon the Borgatti (1990, 1992) software in four important respects: (1) It uses state of the art Bayesian statistical inference, the current standard in computational statistics, (2) It allows items to differ in cultural saliency, (3) It allows informants to differ in guessing biases, (4) It provides a statistical sound Bayesian post predictive test for the single culture assumption of the model. Number (4) is important because it corrects a flaw in the earlier software, where the one-culture assumption was not assessed properly.

A second place where new Bayesian software was provided during the period of the ARO grant was that the early GCM and later models for other response formats (to be described in Section III) were formulated hierarchically, and Bayesian hierarchical inference is developed for them (Papers 8, 14, 16, 17, 23, 24, 25). All papers but Paper 23 and 25 are published or in press, and they include two in *Journal of Mathematical Psychology*, two in *Psychometrika*, and one in *Sociological Methodology*. All these papers provide software for conducting statistical inference for the models.

A third place where Bayesian software has been provided is Paper (22), submitted to *Journal of Statistical Software*. The paper describes "*CCTpack: Cultural Consensus Theory Applications to Data*. R package version 0.99", which is a GUI provided as an R package that provides hierarchical Bayesian inference options for a collection of old and newly developed CCT models. The package requires JAGS (Just Another Gibbs Sampler, Plummer, 2012), which is a freely available package for using Markov Chain Monte Carlo (MCMC) methods to perform Bayesian hierarchical inference on a wide class of statistical models.

III. New CCT Models

Almost all of the work in CCT prior to the last five years concerned the General Condorcet Model (GCM) for dichotomous (True/False) data. While it is easy to conduct a True/False test with groups of informants, the responses to questions in such a design provide only one bit of information. In the period of this grant, with some overlap with an earlier grant from The Air Force Office of Scientific Research to the PI, several new CCT models were invented along with suitable Bayesian hierarchical inference software. These models will be described in the next subsections.

A. Models with a Don't Know Option

There are many survey questionnaires that provide questions with three response options, "True," "False," and "Don't Know" (DK). For example, science questions on the General Social Survey, for example the 2010 General Social Survey (in collaboration with the National Science Foundation) uses a set of 12 questions, 11 of which are True/False format with a `Don't Know' option provided. Evidence suggests that such questionnaires can be problematic because a DK response is often scored as incorrect, and yet different people, for example women on average, have been shown to be more likely than others to use a DK response. In Paper (24), in press in *Sociological Methodology*, the GCM was modified to handle a DK response, and the modification allowed different informants to

have different biases for using this response. Inference software was presented for the new model, and it was applied to the science social survey questions mentioned above as well as Palmore's (1998) "Facts on Aging Quiz". Covariates were also incorporated into the model, and it was shown that a consensus answer key that corresponded to the scientifically correct answers was estimated despite isolated sub-cultures that were discernable through the covariates.

In Paper (25), under review in *American Journal of Psychology*, the model with a DK response was used on some questions about "what constitutes behavior?" The informants were scientists in three societies of behavioral biologists and an additional group of undergraduate psychology majors. The questions posed acts of hypothetical organisms (e.g. a heart pumps blood, a rat runs up a tree), and the informants had to decide using the three response options whether a particular act constituted behavior. The new CCT model was able to find a consensus definition of behavior as well as to score each behavioral or non-behavioral act for its cultural saliency relative to this definition. The paper proposes CCT more generally as an approach to examining whether or not a group of scientists share a consensus on the use of a particular scientific term.

B. Models for Assessing Ties in a Graph

In Paper (20), a model is developed for assessing ties (edges) in a graph, e.g. a social network. In this model, informants respond dichotomously (Present, Absent) to questions about whether or not there is a tie between two nodes in a graph. One earlier CCT paper on graph aggregation was Batchelder, Kumbasar, and Boyd (1997). The main addition to this model provided in Paper (20) is that the researcher can impose constraints on the answer key (consensus graph) that are not assumed to hold for the responses of each informant. In this paper, a two-cluster balance structure (e.g. Cartwright and Harary, 1956) is imposed on the consensus graph. One interesting facet of this work is that the addition of a required structure on the answer key requires that a model-specific MCMC sampler had to be constructed to conduct the Bayesian inference for the model. Several studies were run, for example cities were selected out of two states, and for each pair of cities the informant had to say 'Same State' or 'Different State.' In another study, informants were given two states and asked whether or not they shared a border. In this work, it was realized that tie judgments could be based on either nodal knowledge (e.g. the state where the city is) or tie knowledge (is there a shared border), or both. Paper (20) developed a tie-based model; however, the more difficult model to design and estimate is a nodal-based model. The additional difficulty is due to the fact that if an informant knows the attribute of a particular node, that knowledge is operative for any tie in the graph involving that node. In the later case, a nodal model has recently been worked out along with a model-specific MCMC sampler, but it has not been tested against real data or published yet. Apart from the single published paper mentioned above, this area is currently an active one in out lab.

C. *Models for Ordinal (Likert Data)*

One way to collect more information about a question is to ask for degrees of truth on an

ordered scale. One example of ordinal CCT data would be for informants to rate sentences of a certain type on grammaticality using a seven-point scale, from clearly grammatical to not grammatical. Another example is to rate a particular city in terms of its properties (quality of weather, of nightlife, etc.). In Paper (16), in press in *Psychometrika* (down loadable from their website, but not yet appearing in the journal), a new CCT model for ordinal data is provided, and applied to real data from the above two examples. The model assumes that each item occupies a latent place on a continuum, and that each informant draws an item placement from a Gaussian distribution centered on the item placement and with a variance depending on the competence of the informant and difficulty of the item. Each informant chops the continuum into regions corresponding to the ordinal categories. Different informants can chop differently depending on their biases for extreme or middling ratings as well as ratings shifted up or down the scale. In addition, the new model allows for one to test for and identify two or more consensus answer keys (latent item locations on the continuum) from the data. This addition to CCT models is described in some detail in Section IV to follow. It was found that in the grammaticality data there was a single consensus truth; however, in the cities data the model was able to estimate a separate answer key for each city tested (Irvine, CA; New York City, Miami).

The paper provides a detailed comparison between CCT modeling and IRT (item response theory) modeling popular in psychometric test theory. IRT models assume knowledge of the correct answers to the questions and they attempt to estimate the ability of the test takers, on the other hand, CCT models are interested in estimating the correct answers to the questions as seen by the test-takers, and only secondarily to estimate the ability (cultural competence) of the informants. Apart from these essential differences, both theories use somewhat similar statistical specifications of their models.

D. *Models for Continuous Data*

In Papers (23) and (21), CCT models for continuous type response data are provided. Paper (21) (in press in *Educational and Psychological Measurement*) uses classical Maximum Likelihood Estimation techniques instead of hierarchical Bayesian inference, and the model is applied to a published data set consisting of experts rating 50 essays on 5 attributes. Multiple correspondence analysis was used to convert the five ordinal ratings into a continuous scale. The model allows several kinds of rater biases and item easiness parameters to be estimated. Fourteen raters, including two of the experts, provide the data for the model, and a single consensus answer key was estimated. As expected, the model is able to pick up the experts as the ones with high competence. Paper (19) provides some available software for these continuous models.

In Paper (23) (under review *Journal of Mathematical Psychology*), a hierarchical model for continuous data such a probability judgments is constructed. Its underpinnings are similar to the ordinal model in Paper (16) except bias is represented differently in that it is applied to the latent draws rather than to the cut points of the ordinal model. Bayesian hierarchical inference is developed for the model (see Section II). The model is applied to forecasting the likelihood of future events and to evaluating various eating practices connected with health issues. The model was extended to allow more than one answer key (described in Section IV). It turns out that the model selected a single answer key for the

forecasting questions despite considerable heterogeneity in the responses; however, for the eating-practices-data there were two discernable answer keys, with one indicating considerable concern and compliance with popular health issues.

IV. Tests of the One Culture Assumption

Crucial to the successful use of a CCT model is to statistically evaluate the evidence for a single culture solution. If the single culture assumption is rejected, then either the data support more than one consensus answer keys or there is no discernable consensus in the data. One of the major developments in the period of the grant has been the development of tests for one or more than one answer keys. These developments are seen in Papers (14), (16), (17), and (23). In the proposal for the ARO grant, the following was written:

"There are two cases where CCT can be used in studying multiple cultures: (1) the cultural membership of each informant is obtainable from covariates available on the informants; and (2) the cultural membership of each informant is latent (unobserved). In both cases, an appropriate CCT model would specify two or more latent answer keys as parameters. The second case is the more difficult one because the CCT model must also specify latent group membership parameters. In this case, the model would take the form of a finite mixture model … "

In Paper (8), a major theorem for the GCM was developed that bears directly on the one-culture assumption of this model. The key is to construct a matrix of informant-informant correlations taken over items, say $\mathbf{R} = (r_{ij})_{N \times N}$, where r_{ij} is the Pearson correlation between informants *i* and *j* calculated over their dichotomous responses to the items (from the 2 by 2 table consisting of the number of items that both said 'True," *i* said "True" *j* said "False", etc.). The theorem shows that the correlation matrix has a one-factor structure, that is the off-diagonal terms are approximated by $r_{ij} \approx a_i a_j$, where, for example, a_i is the correlation between informant *i*'s responses and the latent consensus answer key. Such a property is testable by a suitable factor analysis of the correlation matrix followed by an examination of the scree plot of the eigenvalues to see if the data supports a one-factor structure. While checking a scree plot for a single factor structure is generally handled with various controversial 'rules-of-thumb,' Paper (14) developed a graphic Bayesian post-predictive test (see paper for details) for the one answer key assumption. The test worked well to discern when a one-culture assumption was met with simulated data, and it has been incorporated into the software for the GCM model.

Two mathematical developments of the one-factor test allowed tests to be developed for other models and other numbers of cultural truths. The basic theorem for the case of several answer keys in the GCM states that the correlation of two informants, *i* and *j*, over the items is a product of three terms: (1) The correlation of *i* with his or her answer key (like the a_i above); (2) The correlation of *j* with his or her answer key; and (3) The correlation between the two answer keys. The third factor in the above formulation implies that the scree plot will exhibit more than one factor in case there is not a single cultural truth, and it was developed in Paper (14) into a Bayesian post-predictive test for multiple cultural truths. This approach worked well on simulated data, and it revealed

multiple cultures in several well-known anthropological data sets that were not known before the model analysis (see papers).

The results described so far were proved for the GCM, namely the main model for dichotomous data. As a consequence it became important to see if this approach could be applied to the new CCT models described Section III. It turned out that this approach either caries over completely (for example for the continuous model) or holds to a good approximation for the other models. So the approach for detecting the number of cultural truths based on the scree plot of the informant-by-informant correlations has been incorporated into the Bayesian hierarchical software provided in Paper (22). Viewed generally, these developments have extended CCT to a form of substantive, model-based clustering, where informants are clustered into consensus truth patterns based on their responses to the same set of questions.

V. Other Research Supported by the Grant

It turns out that CCT models for categorical data (e.g. the GCM and the model for the don't know response) are examples of Multinomial Processing Tree (MPT) models, which are a major variety of cognitive models first invented by the PI and his students in the 1990s, e.g. Batchelder and Riefer (1999). Several new papers in this area were published, some with the support of this grant. These include Papers (12), (13), and (18). Paper (12) is an invited encyclopedia piece on MPT type models, Paper (13) develops Bayesian hierarchical inference for one of the most popular MPT models, and Paper (18) is a reply in *Psychological Bulletin* to some authors who posed questions about the viability of MPT models. In addition, Paper (15) is a major paper in the area of human problem solving published in *Journal of Problem Solving* with some support from the grant. While it is unrelated to CCT models, it has an interesting set of brainteasers in it, and it was featured in a keynote address by the PI in a Dagstahl Castle Conference on Problem Solving involving formalists in the computer sciences and psychology.

VI. Conclusion

During the period of the grant, it has been possible to greatly expand CCT modeling well beyond its usual application area in cultural anthropology. Papers publishing CCT model analysis have appeared or are under review in anthropology outlets (Papers 10,11), sociology outlets (Papers 20, 24), psychology outlets (Papers 8, 14, 16, 17, 23, 25), education outlets (Papers 19, 21), and statistics outlets (Paper 22). The PIs Cognitive Psychometrics Laboratory has included five undergraduates, five graduate students, one Post Doctoral Fellow, one Sabbatical visitor, and several Professors. Many of these researchers gladly involved themselves with CCT research projects, although many of them received no financial support from the ARO grant. With the new CCT Graphic User Interfaces (Papers 10,11,22), there is hope that researchers who are already using CCT models will greatly improve their statistical inference potential, and newcomers to the models will find them useful for their purposes.

There remain CCT application areas and model developments that are in need of additional research and application. Many of them fall into the areas of aggregating ties in a social or covert network, discerning events that happened from eyewitnesses, and detecting intentional prevaricators among informants. It will be the intention of the PI to apply for more funding from the ARO if the results of the current grant are considered to have a proper level of importance for additional funding.

References

Batchelder, W.H., Kumbasar, E., and Boyd, J.P. (1997). Consensus analysis of three-way social network data. *Journal of Mathematical Sociology*, **22**, 29-58.

Batchelder, W.H. and Riefer, D.M. (1999). Theoretical and empirical review of multinomial processing tree modeling. *Psychonomic Bulletin & Review*, **6**, 57-86

Batchelder, W.H. and Romney, A.K.(1986). The statistical analysis of a general Condorcet model for dichotomous choice situations. In B. Grofman and G. Owen (eds.) *Information Pooling and group decision Making: Proceedings of the Second University of California Irvine Conference on Political Economy* (pp.103-112). Greenwich, Conn.: JAI Press .

Batchelder, W.H. and Romney, A.K. (1988). Test theory without an answer key. *Psychometrika*, **53**, 71-92.

Batchelder, W.H. and Romney, A.K. (1989). New results in test theory without an answer key. In E.E. Roskam (Ed.) *Mathematical Psychology in Progress* (pp.229-248). Heidelberg: Springer-Verlag.

Borgatti, S. P. 1990. Using ANTHROPAC to investigate a cultural domain: Part 3 *CAM Journal* 2(3):10-10.

Borgatti, S. P. 1992. *ANTHROPAC 4.00 user's guide*. Columbia: Analytic Technologies

Cartwright, D., Harary, F.(1956). Structural balance: a generalization of Heider's theory. *Psychological Review*, **63**, 277-293.

Palmore, E. (1998). Facts on aging quiz, second edition. New York: Springer Publishing Co.

Plummer, M. (2012). Rjags: Bayesian graphic models using mcmc. R package version 3.2.0 (http://CRAN.R-project.org/package=rjags).

Karabatsos, G., and Batchelder, W.H. (2003). Markov chain estimation theory methods for test theory without an answer key. *Psychometrika*, **68**, 373-389.

Romney, A.K., Batchelder, W.H. (1999). Cultural consensus theory. In R. Wilson and F. Keil (eds.). *The MIT Encyclopedia of the Cognitive Sciences* (pp.208-209). Cambridge, Mass.: The MIT Press.

Romney, A.K., Batchelder, W.H., and Weller, S.C. (1987). Recent applications of cultural consensus theory. *American Behavioral Sciences*, **31**, 129-149.

Romney, A.K., Weller, S.A., and Batchelder, W.H. (1986). Culture as consensus: A theory of culture and informant accuracy. *American Anthropologist*, **88**, 313-338.

Weller, S.A. (2007). Cultural consensus theory: Applications and frequently asked questions. *Field Methods*, **19**, 339-368.